

Effects of effort and difficulty on human preference for a stimulus: Investigation of the within-trial contrast

Masashi Tsukamoto¹ · Kenichiro Kohara² · Koji Takeuchi²

Published online: 12 September 2016
© Psychonomic Society, Inc. 2016

Abstract The within-trial contrast hypothesis (WTC) provides a more parsimonious explanation for the phenomenon that humans and animals prefer outcomes that follow more effortful events to outcomes that follow less effortful events (Zentall, 2013). We conducted two WTC experiments with human adults. In Experiment 1, we manipulated the difficulty of a preceding event by varying the interresponse time and the limited-hold interval during differential reinforcement with a low response rate schedule, to examine the effect of effort on the preference for the subsequent stimuli. In Experiment 2, we attempted to identify the variables that had affected the results of Experiment 1, by manipulating time as the delay of reinforcement. The results showed preferences based on WTC only when participants made a high rate of incorrect responses in the preceding event, which was used as an index of the strength of individual effort. These results extend the findings of previous human WTC studies and suggest that the difficulty of a task could serve as an aversive event that affects the WTC effect. It is possible that an index based on performance in the preceding event would provide useful information for predicting the contrast effect.

Keywords Human · Preference · Effort · Difficulty · Delay

When humans expend a greater amount of energy to achieve a goal, they generally appreciate the results more than if less

work was necessary. For example, Aronson and Mills (1959) manipulated task difficulty and reported that participants who were required to read some embarrassing sexual descriptions in order to join a discussion group placed greater value on the group as the severity of the required task increased. Recently, a similar phenomenon has been reported in nonhuman animals (Zentall, 2013). A series of studies beginning with Clement, Feltus, Kaiser, and Zentall (2000) has examined whether there are differences in preferences for outcomes that follow two kinds of preceding tasks involving different efforts. On one trial, an effortful task is followed by a presentation of a simultaneous discrimination. On another trial, an easier task is followed by a presentation of different task. After going through training on these tasks, when animals are required to choose between two positive stimuli (S+), they develop a preference for the S+ that followed the more effortful task in training. Earlier researchers had explained the difference in preferences in humans depending on their effort in terms of higher-order psychological process such as cognitive dissonance (e.g., Aronson, 1969). However, the observation of preferences for outcomes that follow more effortful tasks in nonhuman animals suggests that some common factors operate in both humans and nonhuman animals.

Clement et al. (2000) propounded a more parsimonious explanation, which is now called the *within-trial contrast hypothesis* (WTC), to account for the phenomenon in nonhuman animals. According to the WTC (e.g., Zentall, 2010) one can set the relative value of the trial at zero at the beginning of each trial. Then, some effort imposed during the trial results in a negative shift in the value. Finally, because the relatively effortful requirement results in a greater negative shift, there is a greater positive shift in the value than with the other, less effortful requirement, and a preference is produced for the positive stimulus that follows the more effortful response.

✉ Masashi Tsukamoto
10mp011@stu.meisei-u.ac.jp

¹ Department of Psychology, Graduate School of Humanities, Meisei University, Tokyo, Japan

² Department of Psychology, School of Humanities, Meisei University, 2-1-1, Hodokubo, Hino-shi, Tokyo, Japan

The WTC has been tested in a variety of contexts, using an experimental procedure similar to that of Clement et al. (2000), and a certain amount of reproducibility has been demonstrated (e.g., Clement & Zentall, 2002; Friedrich & Zentall, 2004; Gipson, Miller, Alessandri, & Zentall, 2009; Kacelnik & Marsh, 2002; Zentall & Singer, 2007). These studies investigated the validity of the WTC to measure the influences of response frequency (e.g., Clement et al., 2000), delay of reinforcement (DiGian, Friedrich, & Zentall, 2004), and other efforts. Furthermore, the experimental paradigm of Clement et al. (2000) has been used with humans. Studies have shown a significant preference for the stimuli that follow the more effortful task (Alessandri, Darcheville, Delevoeye-Turrell, & Zentall, 2008; Alessandri, Darcheville, & Zentall, 2008; Klein, Bhatt, & Zentall, 2005).

Although the WTC has been demonstrated in numerous studies, other studies have reported failures to find the contrast effect in both humans and nonhuman animals (Arantes & Grace, 2008; Aw, Vasconcelos, & Kacelnik, 2011; Shibasaki & Kawai, 2008; Shibasaki & Kawai, 2011; Vasconcelos, Urcuioli, & Lionello-DeNolf, 2007; Vasconcelos & Urcuioli, 2008). For example, Arantes and Grace (2008), who tried to replicate the contrast effect, trained pigeons using the same response requirement (FR20 vs. FR1) as Clement et al. (2000). They found that although the preference for the following stimuli became greater with increased amounts of training, the direction of the preference was inconsistent with the prediction based on the WTC. These studies have provided evidence about the boundary conditions of the WTC, and also gave rise to an argument about procedural variables (e.g., Zentall, 2013; Zentall & Singer, 2007). As one of the main procedural variables, the preceding tasks were not sufficiently aversive events for the individuals (Zentall, 2013). In Arantes and Grace (2008), they used 20 pecks as an effortful requirement, and their pigeons had prior experience with a variable-interval schedule (VI). It is known that VI is designed to produce a constant rate of response (Ferster & Skinner, 1957). Thus, the 20 pecks that served as an effortful event were less likely to arouse an aversive state for the pigeons that had previously performed on a VI (Zentall, 2013).

Some studies have addressed the issue of how individuals evaluate prior events. Alessandri, Darcheville, Delevoeye-Turrell, and Zentall (2008), who tried to replicate the contrast effect with humans, assessed the participants' preference for tasks that included various combinations of force and time before the training. Then they assigned the most preferred task to the low-effort trial and the least preferred task to the high-effort trial. They succeeded in obtaining the contrast based preference for the stimuli that followed the nonpreferred task in training. This result suggests that if participants are trained using tasks based on their preferences, then the contrast effect will be observed. In contrast, Vasconcelos et al. (2007) reported that they trained pigeons using more effortful response

requirements (FR80 vs. FR1) than the original contrast effect demonstration (Clement et al., 2000). They failed to replicate the contrast effect even though the independent measures suggested that the preceding tasks were aversive for the pigeons. Because these studies that had inferred the aversiveness of tasks showed inconsistent results, it seems that the procedural variables proposed by Zentall (2013) are not critical as explanations for the failure to replicate the contrast effect.

However, the tasks' aversiveness implied in the two previous studies might be separate from the performance of the preceding tasks. Alessandri, Darcheville, Delevoeye-Turrell, and Zentall (2008) presented two tasks that consisted of various combinations of force (high vs. low force) and time (1 vs. 5 s) simultaneously. They assessed the participants' preferences for the tasks before the training by instructing them to choose a task according to their preference. Vasconcelos et al. (2007) measured the response latencies for each preceding task to confirm whether pigeons could anticipate low or high response requirements when they saw the stimulus of the preceding tasks. Vasconcelos et al. observed much longer latencies on the large-FR trials (FR80) than on the small-FR trials (FR1). Although these data address an important aspect related to the aversiveness of the preceding tasks, they were evaluated independently from the performance in the preceding tasks. One was assessed prior to the training, and the other was measured just before the preceding tasks in training. Thus, there was no guarantee that the manipulated effort served as an aversive event for the individuals.

According to the WTC, the contrast effect occurs when the preceding event serves as a relatively aversive event for an individual (Zentall, 2013). Although in most research on the WTC the difference in effort between the two trials has been manipulated using various types of responses and schedules, there seems to be room for consideration regarding the performance in the preceding tasks. For example, there are several kinds of responses: key pecking, traveling between two perches by flying or walking (Kacelnik & Marsh, 2002), mouse click (Klein et al., 2005), and so on. There are also several forms of delay, ranging from active ones—such as responding in fixed-interval schedules of reinforcement (FI)—to passive ones—such as just waiting for the reinforcement. Because some stimulation associated with responses, such as kinesthesia, may serve as a reinforcement (Kish, 1966), this could affect whether the contrast effect is obtained among individuals even with apparently effortful requirements. Further examination focusing on performance in the preceding tasks with manipulations of effort for individuals would be necessary to determine the effects of the variables on the contrast effect.

We conducted two WTC experiments with humans. In the present study, we equalized the numbers of responses to the operandum per trial to minimize the influence of uncertain stimulation, and in Experiment 1 we required differential

mediating behaviors as efforts between trials. In that experiment, we manipulated the difficulty of the preceding tasks as the one variable that affects the effort in the preceding tasks, and calculated the incorrect responses in the preceding tasks as an index of the strength of the effort for the participants. In Experiment 2, we attempted to identify the variables that had affected the results of Experiment 1 by manipulating time as the delay of reinforcement. The purpose of the present study was to determine whether the relative efforts elicited by the two preceding tasks and the absolute effort indicated by the performance in the tasks would affect the contrast effect. Additionally, we also examined the effect of the effort produced by the response requirements on the outcomes in terms of the similarity and difference of the preceding tasks in two experiments.

Experiment 1

In Experiment 1, we focused on the relation between the difficulty of the preceding tasks and that of the outcome with human participants. We used the experimental paradigm of Clement et al. (2000) to examine the effect of the difficulty of the preceding tasks on the outcome. In the present experiment, we manipulated the difficulty of tasks quantitatively, on the basis of the technique of Conrad, Sidman, and Herrnstein (1958). Conrad et al. trained rats using a differential reinforcement of low response rate schedule (DRL). They imposed a limited hold (LH) on the time interval after which a response could be reinforced, to get rats to learn very precise control of the rate of responding. For example, on a DRL-10 LH-2.0, only responses that are emitted between 10 and 12 s after a preceding response will be reinforced; responses emitted less than 10 s or more than 12 s after a preceding response are unreinforced. If a DRL requires a long pause followed by a certain response, then individuals under the DRL sometimes develop mediating behaviors before they emit that response (Laties, Weiss, Clark, & Reynolds, 1965; Zuriff, 1969); especially, human participants may count the time until the reinforcement is available. Furthermore, participants might engage in greater effort than if an LH were not imposed, because the harder the temporal constraint is, the more participants have to count the time carefully. In this experiment, we used two kinds of quasi-DRLs that required participants to respond to the operandum twice per trial, and we manipulated the interresponse time (IRT) and the interval of the LH. In this experiment, we examined whether more effortful tasks, which were affected by an operation involving greater temporal difficulty, produced a preference for the stimuli that followed, by using the rates of incorrect responses on the two preceding tasks to assess the strength of effort for each participant in a preceding task.

Method

Participants

A total of 24 undergraduate and graduate students at a private university in Japan were randomly assigned to two groups: an LH-0.6 group ($n = 12$; six female, six male) and an LH-4.0 group ($n = 12$; five female, seven male). The data from four additional participants who said in the postexperiment interviews that they had little or no difficulty in the “difficult tasks” were not included because similar participants had yielded inconsistent data in our preliminary experiments. The data from another four students who could not meet the criterion in the training were also excluded.

All of the participants majored in psychology but were not familiar with the topic of this study. Each participant was fully debriefed at the end of the experiment and was given a \$5 gift certificate.

Apparatus

All participants were trained and tested with a program created with Microsoft Visual Basic 2010, run on a 14-in. notebook computer (ThinkPad Edge E420, Lenovo).

Materials

A white circle drawn with a vertical line and a white circle drawn with a horizontal line served as the discriminative stimuli in the initial tasks. Before starting this experiment, participants were shown 14 free-form line shapes drawn with Microsoft Paint and were instructed to select three they preferred and three they hated (Fig. 1). The eight unchosen stimuli were used as the discriminative stimuli in the terminal tasks. Four of the stimuli were assigned to the group of stimuli that followed the more difficult tasks, and the remaining four stimuli were assigned to the group of stimuli that followed the less difficult tasks.

Procedure

Pretraining This phase was set up for participants to learn the operation that was required in the next phase (training). In this experiment, two kinds of “timing behaviors” were used for the main tasks. On half of the trials, when a vertical line was drawn at the center of the screen under the presentation of the circle, each participant pressed the space bar that made a start of timing, and after 10 s they pressed the space bar again and got feedback (hereafter referred to as the *IRT-10 task*). The remaining trials employed the same task, except a horizontal line was drawn at the center of the screen under the presentation of the circle, and the interval of time between the two

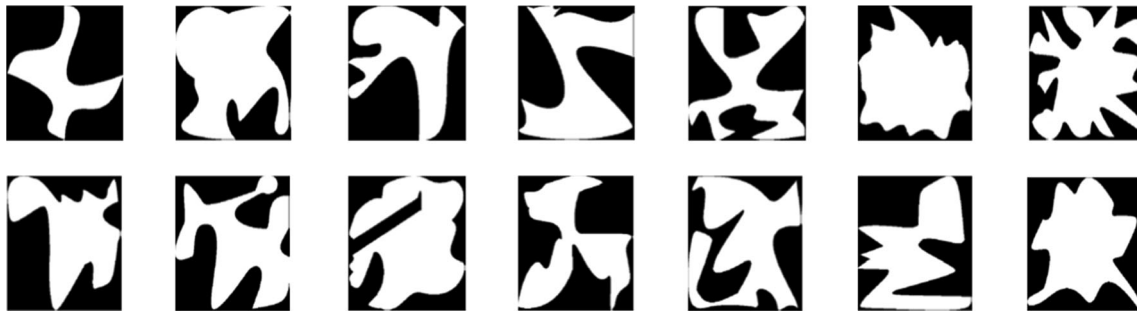


Fig. 1 Positive stimuli and negative stimuli (S+ and S−) used in the terminal tasks in each experiment

space-bar pressings was 2 s (hereafter referred to as the *IRT-2 task*). During the trials, no cues showed the time.

To manipulate the difficulty of the tasks, temporal constraints (Conrad et al., 1958) were imposed on all trials. The second response was reinforced within a brief period of time, around 10 s on the IRT-10 task and around 2 s on the IRT-2 task. Half of the participants who were assigned to the LH-0.6 group could emit the second response from 9.7 to 10.3 s after the first response on the IRT-10 task, and from 1.7 to 2.3 s after the first response on the IRT-2 task. The remaining participants, who were assigned to the LH-4.0 group, could emit the second response from 8 to 12 s after the first response on the IRT-10 task, and from 0 to 4 s after the first response on the IRT-2 task. Through the use of these limited holds, the IRT-10 task was made relatively more difficult than the IRT-2 task in the LH-0.6 group, whereas there might have been no difference in the degrees of difficulty for participants between the tasks in the LH-4.0 group. When participants emitted the second response successfully, they got the feedback “Correct.” If participants did not emit the second response in the available time, they got the feedback “Too Early” or “Too Late,” and they moved on to the next trial. Pretraining consisted of blocks of 16 trials (eight IRT-10 tasks, eight IRT-2 tasks) and continued until participants got 13 correct feedbacks out of 16 consecutive trials (accuracy rate about 80 %).

Participants were given the following instructions:

There are two kinds of tasks which require you to press the space bar twice. When you see a vertical line, press the space bar once, and then press the space bar again after just 10 s has elapsed since your first response. When you see a horizontal line, you should respond the same as to the vertical line, except the elapsed time between the two space bar pressings should be just 2 s. Although the second pressing response has only a small window of time, around 10 s or 2 s, you should press the space bar at the given time as best you can. You will receive feedback about your response.

Training In this phase, participants performed on two-component chain schedules at random. Each component consisted of two kinds of tasks: the initial task and the terminal task (see Fig. 2). At the onset of trials, participants had to press the space bar one time after the presentation of the word “Ready,” which was shown at the center of the screen, and then performed either the IRT-10 task or the IRT-2 task, which served as an initial task. These tasks also had the LHs that had been used in the pretraining. Participants who were in the LH-0.6 group got the initial tasks with more stringent time limits for correct response times, and participants who were in the LH-4.0 group got the initial tasks with a greater time window for a correct response, as we described above. When participants completed the initial task that they had learned in the previous phase, they moved on to the terminal task. For the terminal tasks, a pair of line-drawn shapes (S+ and S−; see Fig. 1) were presented, with one on the left and one on the right side of the screen, and participants were required to choose between S+ and S− by pressing either the “F” key or the “J” key. The choice of one stimulus (S+) was followed by the feedback “Correct,” and the choice of the other (S−) was followed by the feedback “Wrong.” After this feedback, there was an intertrial interval (ITI) of 1 s, and then participants moved on to the next trial.

For the terminal tasks, because two kinds of trials were initiated with the different IRT tasks, and each of them had two S+s and two S−s, the total number of trial combinations was eight. Furthermore, the positions of S+ and S− (left or right) were counterbalanced. Therefore, training consisted of blocks of 16 trials and continued until participants got 13 correct feedbacks out of 16 consecutive trials (accuracy rate about 80 %). After the criterion was met, further training trials were conducted for some participants, to control for stimulus exposure between the two kinds of trials. Moreover, after reaching the criterion, further training trials were added for the participants who were assigned to the LH-4.0 group until they had dealt with each task with almost the same frequency as the LH-0.6 group, because participants in the LH-4.0 group tended to reach the criterion in fewer trials than did those in the LH-0.6 group. When participants completed the training, they had about a minute’s rest and then moved on to the next phase.

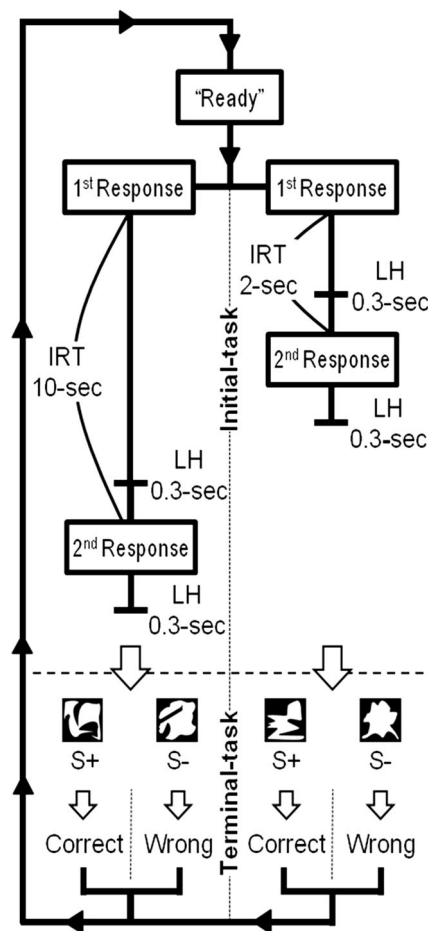


Fig. 2 Summary of the design of training for the LH-0.6 group in Experiment 1. Each response was a space bar pressing. IRT is the interresponse time, and LH is a limited hold. The LH was 0.6 s on each trial in the LH-0.6 group, and 4.0 s on each trial in the LH-4.0 group. The shapes of S+ and S− were pseudorandomly assigned for participants. For the terminal tasks, because two kinds of trials were initiated with the different IRT tasks, and because each of them had two S+s and two S−s, the total number of actual trial combinations was eight. The positions of S+ and S− (left or right) were counterbalanced. A second response that occurred outside the LH was not reinforced and was counted as an incorrect response

Participants were given the following instructions:

In this phase, each trial begins with the presentation of the word “Ready,” which leads to different IRT tasks by pressing the space bar once. If you see the vertical or horizontal line, you should respond as you had practiced earlier. In this phase, when you complete the initial task, two shapes are presented simultaneously instead of the word “Correct.” If you choose the correct shape, you will receive the feedback “Correct.” If you choose the wrong shape, you will receive the feedback “Wrong.” There are several shapes in the terminal tasks. Correct shapes always serve as a correct shape. Wrong shapes also serve as a wrong shape the whole time. You should judge the correct shapes by the form. The training phase

will not be over if you make errors in the initial tasks or choose the wrong shapes in the terminal tasks. Please complete each task without errors as best you can.

Testing In this phase, with the initiation of each trial, a pair of the stimuli from training were presented simultaneously, one on the left and one on the right side of the screen, and participants were told to choose between two stimuli by pressing either the “F” key or the “J” key. Testing consisted of 32 trials: eight trials with the presentation of the two S+s, eight trials with the presentation of the two S−s, eight trials with the presentation of the S+ that followed the IRT-10 task and the S− that followed the IRT-2 task, and the remaining eight trials with the S+ that followed the IRT-2 task and the S− that followed the IRT-10 task. The trials of S+ versus S− items were intended to conceal the purpose of this experiment from the participants. Therefore, these trials were excluded from the data analysis. In this phase, there was no feedback; thus, whatever they chose, the participants would move on to the next trial after an ITI of 3 s. The sequence of trials was randomized, and the numbers of the presentation times and locations of the two stimuli were counterbalanced.

Results

Pretraining

First, we divided the total number of incorrect responses by the total number of trials with respect to each task, to get the rate of incorrect responses in pretraining. Table 1 shows the mean rates of incorrect responses in pretraining. Participants in the LH-0.6 group showed a higher rate of incorrect responses on the IRT-10 task ($M = .48$, $SEM = .05$) than on the IRT-2 task ($M = .22$, $SEM = .04$), whereas participants in the LH-4.0 group showed similar low rates of incorrect responses on both tasks ($M = .06$, $SEM = .03$, vs. $M = 0$, $SEM = 0$, respectively). A 2×2 mixed analysis of variance (ANOVA) was conducted to examine whether there was any differences in the degrees of difficulty between the IRT-10 and IRT-2 tasks in pretraining. The analysis indicated significant main effects of group and trial [$F(1, 22) = 44.58$, $p < .001$, and

Table 1 Mean rates of incorrect responses in pretraining

LH	0.6 s		4.0 s	
	10 s	2 s	10 s	2 s
<i>M</i>	.48 (.05)	.22 (.04)	.06 (.03)	.00 (.00)

M is the mean rate of incorrect responses in pretraining. An incorrect response is a second response that occurred outside of the LH on each task in pretraining. The parentheses in the table represent the standard errors of the means

$F(1, 22) = 30.81, p < .001$, respectively], as well as a significant interaction [$F(1, 22) = 11.01, p = .003$]. To probe this interaction, we examined simple main effects. The analysis revealed significant simple main effects between the mean rates of incorrect responses of the LH-0.6 group and the LH-4.0 group on the IRT-10 task [$F(1, 44) = 55.22, p < .001$], and between the mean rates of incorrect responses of the LH-0.6 group and the LH-4.0 group on the IRT-2 task [$F(1, 44) = 15.99, p < .001$]. Also we observed significant a simple main effect between the mean rates of incorrect responses on the IRT-10 task and the IRT-2 task for the LH-0.6 group [$F(1, 22) = 39.33, p < .001$]. The simple main effect was not significant for the difference between the mean rate of incorrect responses on the IRT-10 task and the mean rate of incorrect responses on the IRT-2 task for the LH-4.0 group [$F(1, 22) = 2.49, p = .129$, n.s.]. These results suggest that the IRT-10 task was more difficult than the IRT-2 task for participants in the LH-0.6 group, whereas there was no difference between the two tasks for participants in the LH-4.0 group. These results also suggest that the shorter LH was more difficult for participants than the longer LH.

Training

Table 2 shows the mean rates of incorrect responses in training. The rate of incorrect responses was calculated in the same manner as in pretraining. Participants in the LH-0.6 group showed a higher rate of incorrect responses on the IRT-10 task ($M = .41, SEM = .02$) than on the IRT-2 task ($M = .16, SEM = .03$). In contrast, participants in the LH-4.0 group showed similar low rates of incorrect responses on both tasks ($M = .05, SEM = .01$, vs. $M = .01, SEM = .00$). A 2×2 mixed ANOVA was conducted to examine whether there was any difference in the degrees of difficulty between the IRT-10 task and the IRT-2 task in training. The analysis indicated significant main effects of group and trial [$F(1, 22) = 140.70, p < .001$, and $F(1, 22) = 41.54, p < .001$, respectively], as well as a significant interaction [$F(1, 22) = 20.60, p < .001$]. To probe this interaction, we examined simple main effects. The analyses revealed significant simple main effects between the mean rates of incorrect responses for the LH-0.6 group and the LH-4.0 group on the IRT-10 task [$F(1, 44) = 129.61, p < .001$], for

Table 2 Mean rates of incorrect responses in training

LH	0.6 s		4.0 s	
	10 s	2 s	10 s	2 s
<i>M</i>	.41 (.02)	.16 (.03)	.05 (.01)	.01 (.00)

M is the mean rate of incorrect responses in training. An incorrect response is a second response that occurred outside of the LH on each initial task in training. The parentheses in the table represent the standard errors of the means

the LH-0.6 group and the LH-4.0 group on the IRT-2 task [$F(1, 44) = 22.26, p < .001$], and for the IRT-10 task and the IRT-2 task in the LH-0.6 group [$F(1, 22) = 60.32, p < .001$]. The difference was not significant for the IRT-10 task and the IRT-2 task in the LH-4.0 group [$F(1, 22) = 1.82, p = .191$, n.s.]. These results confirm that the difficulty of each task that was observed in pretraining remained for participants in the following training.

For the LH-0.6 group, the mean number of trials in training was 97.4 ($SEM = 8.72$). The mean number of initial tasks completely performed by the participants was 68.0 ($SEM = 4.87$). For the LH-4.0 group, the mean number of trials in training was 72.8 ($SEM = 2.18$). The mean number of initial tasks completely performed by participants was 71.0 ($SEM = 2.39$). These data indicate that the mean numbers of presentation times of terminal stimuli and the times per trial between the two groups were almost the same.

We checked the actual IRT from the first pressing response to the second pressing response on each task in the LH-4.0 group because the IRT-2 task in the LH-4.0 group did not require any temporal spacing to meet the contingencies. For the LH-4.0 group, the mean IRTs on the IRT-10 tasks and the IRT-2 tasks that were performed completely by the participants were 9.95 s ($SEM = 0.28$) and 2.06 s ($SEM = 0.60$), respectively. The minimum mean IRT on the IRT-2 task shown by a participant was 1.41 s, and the maximum shown by another participant was 3.37 s. These data indicate that all participants in the LH-4.0 group responded in the same manner as the LH-0.6 group, not with continuous response as in an FR 2.

Testing

The results of testing are shown in Fig. 3. The *x*-axis represents the pairs of stimuli in testing trials ($S+_{10s}$ vs. $S+_{2s}$, $S-_{10s}$

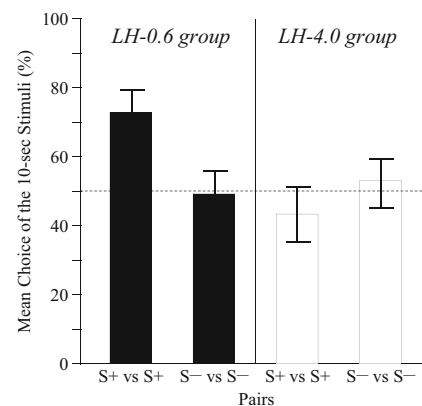


Fig. 3 Average percentages of choices ($\pm SEMs$) of the 10-s stimuli on testing in Experiment 1. The black bars represent the mean choices of the 10-s stimuli in the LH-0.6 group. The white bars represent the mean choices of the 10-s stimuli in the LH-4.0 group. The dotted horizontal line indicates chance level

vs. S_{-2s}). The y -axis represents the mean choices of the stimuli that followed the IRT-10 task in training. To examine whether the mean choice of S_{+10s} in testing was significantly different from chance (50 %), single-sample t tests were conducted. The data was arcsine-transformed before analysis. For the LH-0.6 group, the analysis indicated that choices of S_{+10s} (72 %, $SEM = 7$ %) were significantly different from chance [$t(11) = 3.12, p = .012, d = 0.92$]. For the LH-4.0 group, however, the analysis indicated that choices of S_{+10s} (43 %, $SEM = 8$ %) were not significantly different from chance [$t(11) = 1.00, p = .337, n.s.$]. We also examined whether the mean choices of S_{-10s} (LH-0.6 group: 49 %, $SEM = 7$ %; LH-4.0 group: 52 %, $SEM = 7$ %) in testing were significantly different from chance (50 %). The analysis indicated no significant differences for either the LH-0.6 group [$t(11) = 0.13, p = .901, n.s.$] or the LH-4.0 group [$t(11) = 0.33, p = .746, n.s.$].

In the postexperiment interviews, no participants had been aware of the contingency between the initial task and the terminal stimuli. All participants answered that they had counted the time to complete each task rather than depending on their intuition. At first, some participants counted out vocally, but finally every participant counted the time in the softest of whispers or silently. All participants assigned to the LH-0.6 group reported that the IRT-10 task was more difficult than the IRT-2 task. In both groups, some participants made up names for the S_{+} and the S_{-} in order to learn the roles of the stimuli.

Discussion

In Experiment 1, participants who performed more difficult operations showed a preference for the positive stimuli that followed the effortful trial in training. In contrast to previous studies, in this experiment the numbers of topographically identical responses that were required per trial were equal, whereas the numbers of mediating behaviors that participants engaged in were different between the two types of trials. A temporal constraint was imposed in each task, and that affected the performance of participants, as shown by the rates of incorrect responses. It is possible that such an operation of difficulty had an influence on the strength of effort, unlike the influence of the frequency or force of responses. This was likely to lead a higher rate of preference when we excluded the participants who found no difficulty in the “difficult” task, to ensure the aversiveness of the IRT-10 task. The results of Experiment 1 suggest that a sufficient strength of effort among individuals, even if it was covert, could produce preferences based on the predictions of the WTC.

In contrast to the preference for the S_{+} , we observed no preference for the S_{-} in any group. Clement et al. (2000) reported that pigeons preferred not only the S_{+} that followed greater effort, but also the S_{-} , more than the stimuli that in training had followed less effortful events. This S_{-} effect has

been often discussed in terms of values that transferred from the S_{+} to the S_{-} in each terminal task (Clement & Zentall, 2002). In subsequent studies in humans, however, no significant preference for the S_{-} emerged in test trials that required a choice between the two S_{-} s (Alessandri, Darcheville, & Zentall, 2008; Klein et al., 2005; Shibasaki & Kawai, 2008). In our experiment, some participants applied verbal labels to the S_{+} and the S_{-} , and this might have prompted discrimination between the S_{+} and S_{-} that inhibited the transfer of value from S_{+} to S_{-} (cf. Klein et al., 2005). Furthermore, we used an instruction that encouraged participants to focus on the S_{+} to get them to meet the criterion in training. Thus, the S_{-} may have elicited insufficient interest from participants as a result of this instruction.

We observed no preferences in the LH-4.0 group. For LH-4.0 group, the IRT-10 task was also more difficult than the IRT-2 task, because it imposed more counting and a longer time of concentration for participants. The numerical difference between the two tasks was supposed to produce a preference for the stimulus that followed the longer task, if this difference was sufficient. When compared to the LH-0.6 group, the mean rates of incorrect responses between the two tasks in the LH-4.0 group were almost the same. Thus, participants in the LH-4.0 group could have experienced insufficient negative changes in value during any trials, due to the lack of adequate strength of their efforts. This indicates that it is important to confirm whether the effortful events actually served as an aversive event for the participants, not merely by differentiating the two tasks in effort, but also by using an index related to the performance in the preceding tasks. Previous studies have investigated the aversiveness of the preceding tasks by using some independent measures (Alessandri, Darcheville, Delevoeye-Turrell, & Zentall, 2008; Vasconcelos et al., 2007). Our results extend these attempts and suggest that an index based on the performance in the preceding tasks would serve as a useful reference for prediction of the contrast effect.

Some studies have examined the relation between the delay of reinforcement in prior events and the following stimuli (Alessandri, Darcheville, Delevoeye-Turrell, & Zentall, 2008; Alessandri, Darcheville, & Zentall, 2008; DiGian et al., 2004). These studies reported that when participants were given a choice between the S_{+} that followed a delay and the S_{+} that followed no delay, they showed a significant preference for the delayed as compared to the not-delayed S_{+} . According to these studies, our results could also be interpreted in terms of the effect of time rather than the effect of differential efforts or difficulty, although in a previous report two preprogrammed delays (fixed intervals; FI 3 s vs. FI 18 s) did not produce a preference for the subsequent stimuli (Aw et al., 2011).

Experiment 2

We next explored the effect of delay of reinforcement when the difficulties of the two initial tasks were equal. We also investigated whether the results of Experiment 1 are attributable to the effort that was affected by the difficulty of the prior tasks or by the delay of reinforcement. To clarify this attribution, we adjusted the times per trial in Experiment 2 to almost the same times as in Experiment 1 (i.e., 10 and 2 s) by always using the IRT-2 task as the initial task and adding an 8-s delay to half of the trials. Furthermore, we assigned two kinds of delays to two groups: One was a delay of reinforcement, in which the delay was between an initial task and a terminal task (delay of reinforcement vs. no delay); the other was delay of initiation, in which the delay was between the starting point of the trial and the initial task (delay of initiation of initial task vs. no delay). Because the positive shift of value that occurs at the end of the delay was not paired with S+ in the delay-of-initiation group, the WTC predicts that no significant preference would emerge.

Method

Participants, apparatus, and materials

A total of 24 Japanese undergraduate and graduate students who had not participated in Experiment 1 were randomly assigned to two groups: a delay-of-reinforcement group ($n = 12$; eight female, four male) or a delay-of-initiation group ($n = 12$; seven female, five male). The other features, apparatus, and materials were the same as in Experiment 1.

Procedure

Pretraining This phase was arranged for learning the operation that was required in the next phase (training). The general procedures were the same as those in Experiment 1's pretraining, except that all of the tasks that participants performed were IRT-2 tasks. The instructions differed from those in Experiment 1, in that now in pretraining the participants needed to respond with the IRT-2 task, regardless of the angle of the presented line.

Training The procedure used in Experiment 2's training was similar to that in Experiment 1 (see Fig. 4). On each trial, participants were required to engage in the IRT-2 task. LH-0.6 (0.6 s) was also imposed on every trial. In the delay-of-reinforcement group, a delay was inserted between the initial task and the terminal task. On half of the trials, a correct response to an initial task was followed immediately by a terminal task. On the remaining trials, a correct response on an initial task was followed by a delay of 8 s, and then the terminal task was presented. In the delay-of-initiation group,

there was a delay between the start of the trial and the initial task. On half of the trials, a response following the presentation of the word "Ready" was followed by a delay of 8 s, and then the initial task was presented. On the remaining trials, a response to "Ready" was followed immediately by the initial task. In each group, nothing was displayed on the monitor during the delay. If participants emitted some response during the delay, they were returned to the starting point of the trial with the presentation of "Wrong" feedback.

The instructions were different from those in Experiment 1's training:

Whenever you see the lined shapes, press the space bar once, and then press the space bar again after just 2 s has elapsed from your first response. Some trials have a waiting time. If the monitor blacks out, just wait without responding until the monitor switches on, and keep your eyes on the monitor to engage in the following task immediately.

Testing Participants were required to choose between two S+s (S+_{delay} vs. S+_{nodelay}) or two S-s (S-_{delay} vs. S-_{nodelay}). The general procedure was the same as in Experiment 1.

Results

Training

In the delay-of-reinforcement group, the mean number of trials in training was 79.9 ($SEM = 2.94$). The mean number of initial tasks that were completely performed by participants was 68.0 ($SEM = 0.00$). In the delay-of-initiation group, the mean number of trials in training was 78.3 ($SEM = 2.15$). The mean number of initial tasks that were completely performed by participants was 69.0 ($SEM = 0.37$). These data indicate that the mean numbers of presentation times of the terminal stimuli and the times per trial between the two groups were almost the same.

Testing

The results of testing are shown in Fig. 5. The x -axis represents the pairs of stimuli in the testing trials (S+_{delay} vs. S+_{nodelay}, S-_{delay} vs. S-_{nodelay}). The y -axis represents the mean choices of the delayed stimuli. To examine whether the mean choice of the delayed stimuli in testing was significantly different from chance (50%), single-sample t tests were conducted. The data were arcsine-transformed before the analysis. In the delay-of-reinforcement group, the analysis indicated that choices of S+_{delay} (45%, $SEM = 8\%$) and S-_{delay} (42%, $SEM = 8\%$) were not significantly different from chance [S+_{delay}: $t(11) = 0.40$, $p = .695$, n.s.; S-_{delay}: $t(11) = 0.82$, $p = .432$, n.s.].

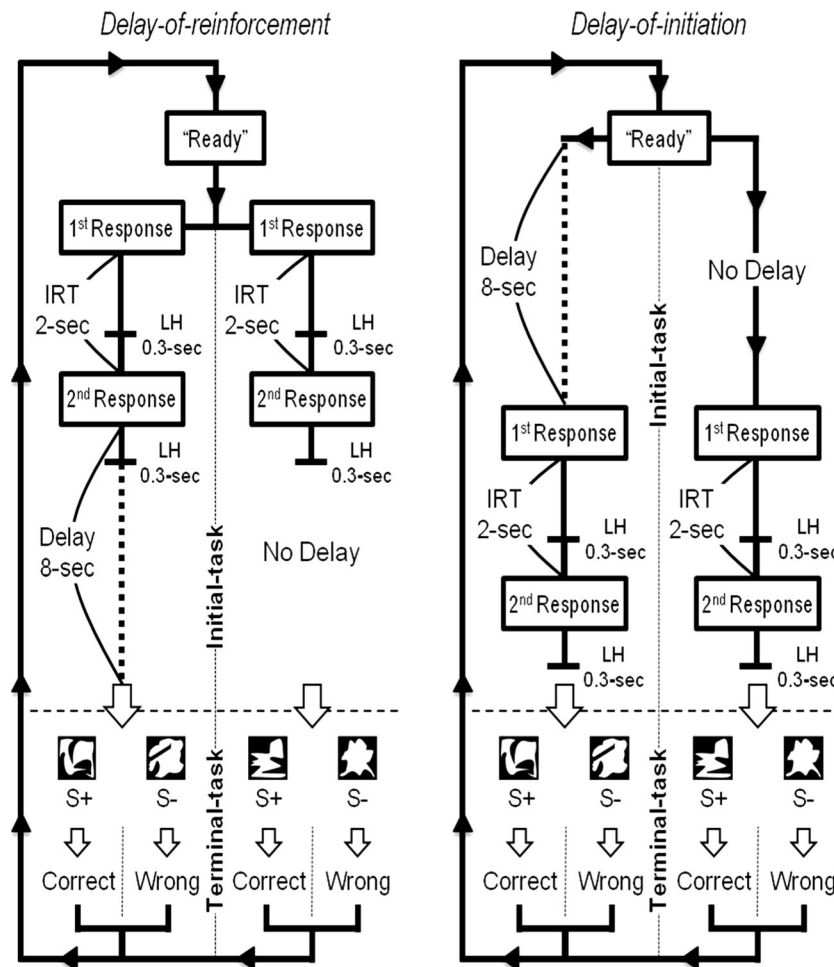


Fig. 4 Summary of the designs of training for the delay-of-reinforcement group and the delay-of-initiation group in Experiment 2. All of the limited holds (LHs) were 0.6 s in this experiment. The shapes of S+ and S− were pseudorandomly assigned per participants. For the terminal tasks,

because two kinds of trials were initiated with the different IRT tasks, and because each of them had two S+s and two S−s, the total number of actual trial combinations was eight. The positions of S+ and S− (left or right) were counterbalanced

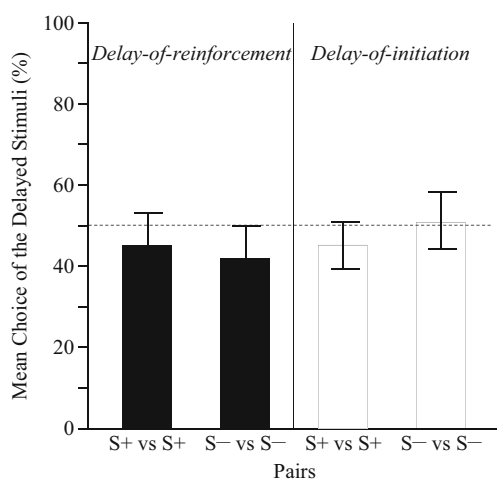


Fig. 5 Average percentages of choices (\pm SEMs) of the delayed stimuli on testing in Experiment 2. The black bars represent the mean choices of the delayed stimuli for the delay-of-reinforcement group. The white bars represent the mean choices of the delayed stimuli for the delay-of-initiation group. The dotted horizontal line indicates chance level

In the delay-of-initiation group, the analysis also indicated that choices of S+_{delay} (45 %, SEM = 6 %) and S−_{delay} (51 %, SEM = 7 %) were not significantly different from chance [S+_{delay}: $t(11) = 0.75, p = .471, n.s.$; S−_{delay}: $t(11) = 0.06, p = .952, n.s.$].

In the postexperiment interviews, no participants had been aware of the contingency between the initial task and the terminal stimuli.

Discussion

In this experiment, we examined whether a delay affects the preference for outcomes, and observed no preferences. This result shows that a delay, known to be a relatively aversive event, does not always produce a preference for the following stimuli. Although Experiments 1 and 2 both required participants to spend almost the same amount of time per trial, the presence or absence of the additional time (8 or 0 s) in Experiment 2 was independent of whether the participants successfully completed the initial tasks. This is because the

programmed time for the prior events was insufficient to affect the preferences for the following stimuli. We argue that the contrast effect found in Experiment 1 resulted not from the time as a delay of reinforcement, but rather from the effort, which was affected by the difficulty of the prior tasks.

One of the reasons that we failed to obtain a contrast effect in this experiment, especially in the delay-of-reinforcement group, which had experimental procedures similar to those in previous studies, could be attributed to our participants. In a previous study that found a contrast effect associated with a delay, the participants were young children, ages 7–8 years (Alessandri, Darcheville, & Zentall, 2008). In general, children are known to be more impulsive than adults; thus, the same delay (8 s in these experiments) would be more aversive to children than to adults.

We also observed no preference for the stimuli that were used in the delay-of-initiation group, but the results were within the scope of the prediction nonetheless. According to the WTC explanation, it is necessary that the positive shift that the participants feel at the end of the initial task be associated with the following stimuli to produce a preference for these stimuli. In the delay-of-initiation group in this experiment, when the delay of 8 s was over, the participants encountered not the discriminative stimuli that were used in the terminal task (a S+/S− pair), but the discriminative stimulus that was used in the initial task (a white circle drawn with a certain line). As a matter of course, although the preference for the stimuli depends on the degree of aversiveness of the prior event, what is important is whether the value changes are associated with the tested stimuli.

General discussion

We conducted two experiments to examine the variables that produce the WTC-based contrast effect. In the first experiment, we required a choice between two stimuli that followed two tasks that differed in effort and difficulty. We found a preference for the stimuli that were preceded by an effortful task in the difficult condition. An additional experiment focusing on the effect of the time delay interval was conducted to determine the variables that affected the contrast effect in Experiment 1. The results of Experiment 2 were inconsistent with those of previous studies: We observed no preference for the stimuli that followed a long delay. These results extend the findings of WTC studies in human beings and show that the difficulty of the task could affect the effort and the preference for the outcomes that followed. They also imply that an index based on the participant's performance in the preceding tasks would provide useful information to predict the contrast effect.

In a series of previous studies in which the delay was manipulated, Alessandri, Darcheville, and Zentall (2008)

obtained the contrast effect in humans, whereas Aw et al. (2011) failed to replicate the effect in starlings. Aw et al. argued, in terms of the nature of effort, that if the main factor affecting the birds is an energetic one, then waiting a certain time would not be sufficient for the birds to manifest the contrast effect. The results of the present study extend Aw et al.'s discussion, and suggest that the failure to find the contrast effect can be attributed not only to the nature but also to the magnitude of the effort. That is, it was possible that the studies that failed to replicate the contrast effect did not induce a sufficiently aversive state for individuals. In our study, we obtained the contrast effect only when we observed a difference in the index that seemed to indicate a greater effort. The present study suggests that the WTC is valid to account for the preference for outcomes that followed effortful events, as long as the prior events are manipulated and designed based on the individual's performance.

In human adults, when we manipulated the difficulty of tasks as effort, we obtained the contrast effect, whereas we observed no preference when there was simply a delay of reinforcement. The responses we required in each task were pressing a space bar and accurate counting of time, which was relatively static and less energetic. The effort that was associated with such responses, however, would have different natures due to the imposed requirements. The differences in the nature of efforts including the effect of difficulty or time between the two experiments' tasks can be divided into two aspects.

The first aspect includes the presence or absence of negative feedback for each initial task and the aversiveness that results from each presentation of negative feedback. As all participants mentioned in the postexperiment interviews, each type of trial of Experiment 1 had not only differential efforts but also differences in the rate of incorrect responses. Because an unreinforced response (incorrect response) under a temporal constraint postpones the reinforcement in a situation in which it had previously been reinforced, this would serve as an aversive event for participants (Amsel, 1958; Melges & Poppen, 1976). In our study, participants tended to make a great number of incorrect responses in Experiment 1, in which we obtained the contrast effect, whereas they seldom made incorrect responses in Experiment 2, in which we failed to replicate the contrast effect. In the LH-0.6 group in Experiment 1, it is possible that negative feedback due to participants' errors postponed some reinforcements (positive feedback or approximation to the end of the experimental situation). Then the negative feedback established correct feedback (presentation of the terminal task) as a strong reinforcer, while difficult tasks that were associated with incorrect responses became more aversive. Our results suggest that activities that can be completed but that involve a certain level of difficulty (but not so difficult as to lead to dropping out) are likely to enhance the value of the reinforcer. In terms of the

difficulty of tasks, a similar change of value has been reported in classical research (Aronson & Mills, 1959). In the present experiment, we found that the positive change occurred using a different kind of paradigm. All participants, however, were not aware of the relation that each pair of S+ and S− followed each initial task; thus, it is not likely that verbal reasoning affected participants' attitudes.

The second aspect includes a difference in the ways of approaching the required efforts. In the case of using the delay as a component of effort, participants were just waiting for the presentation of a terminal task on each trial, so the strength of the response requirement was relatively weak. On the other hand, when we manipulated the difficulty of the tasks, there was a strong response requirement due to the necessity of accurate timing in the initial tasks. That is, the efforts in Experiment 1 that obtained the contrast effect required the participants to respond constantly. Aw et al. (2011) reported that starlings did not show a preference for the outcomes when the initial task consisted of an FI, which contained a waiting time, until a reinforcement was set up, whereas they did show the preference for the outcomes that followed energetically expensive responses, such as flying between the travel perches at the far ends of the cage. In general, an FI has a low rate of responding correlated with nonreinforcement at and near the beginning of the interval (Ferster & Skinner, 1957). When compared to ratio schedules, which require a constant rate of responding, the strength of the response requirements is weak in interval schedules, and this might result in a failure to replicate the contrast effect. An increased response requirement, especially an effort that requires some responding constantly, seems likely to be much more aversive for individuals.

In this study, although we have discussed some possibilities about how the qualitative differences between preceding efforts affect the preferences for the following stimuli, we have not yet provided a comprehensive explanation of the contrast effect. We will need to focus on responses that involve different efforts.

Another explanation may plausibly interpret the results of our Experiment 1. Some possibilities overlap in the mechanism underlying the relationship between the negative feedback and the contrast effect. One is the aversiveness that results from the uncertainty of the tasks. Especially in Experiment 1, because we selectively reinforced accurately counting a fraction of a second, participants could not be sure that they would succeed on the current trial. Uncertainty occurred through repeatedly experiencing failure in the initial tasks, and then the initial tasks associated with the uncertainty might gradually have become aversive events for participants, rather than the aversive response to each negative feedback. Regarding this point, Clement and Zentall (2002), who manipulated the probability of reinforcement in the terminal tasks (high vs. low probability), reported that they observed a preference for the terminal stimuli when pigeons had anticipated a

low probability of reinforcement during the initial tasks. Although our study is different from that of Clement and Zentall, in that the presence or absence of reinforcement depended on the participants' performance, further research will be necessary to determine the effect of uncertainty in the initial tasks on the contrast effect.

We prepared not only differential IRTs, but also different LHs, to emphasize the importance of manipulating effort with sufficient strength for each individual. However, as we discussed, given that the negative feedback rather than the effort functioned as the aversive events that produced the contrast effect, we should use tasks that have the same IRT but different LHs (i.e., a hard vs. an easy LH), in order to examine the effect of the difficulty of the tasks more rigorously. If negative feedback rather than effort could affect the contrast effect in that situation, then we should also obtain the preference for the conditioned reinforcer that followed the hard LH. However, it is also necessary to consider the length of the IRT. Although the 10-s IRT that we used in Experiment 1 could be expected to produce the contrast effect (e.g., an IRT-10 task with LH-0.6 vs. an IRT-10 task with LH-4.0), we also confirmed a higher rate of incorrect responses on the IRT-2 task in the LH-0.6 group than in the LH-4.0 group. If further studies cannot obtain the contrast effect using the same IRT-2 task with different LHs, we will need to reconsider our findings in terms of the interaction between LH (difficulty) and IRT (time).

In the present study, we focused on the difficulty of tasks and attempted to replicate the contrast effect with humans. Is it possible that the difficulty of tasks could produce the contrast effect with animals? The tasks that we used were types of DRL. It is known that animals under the DRL sometimes show mediating behaviors such as pacing (Ferster & Skinner, 1957), tail nibbling (Laties et al., 1965), and so on. If such behaviors can serve as efforts, then a situation similar to that of Experiment 1 might produce preferences based on the WTC in animals. On the other hand, Melges and Poppen (1976) reported that monkeys who face the lack of, or a delay of, expected rewards on the DRL show emotional behaviors such as “bursts,” “activity,” and “vocalization.” Given that animals also experience a state of aversiveness under the DRL, it is possible that a contrast effect that is affected by the difficulty of tasks would occur in animals. In the present study, however, we basically shaped the required responses with our verbal instructions. It is unclear how we could reconstruct a similar situation for animals.

Finally, more than one cognitive factor might affect the positive contrast effects in humans (Zentall, 2013). In our study, no participants were aware of the relationship between the preceding effort and the following outcomes. However, we should consider adopting an appropriate procedure to examine the WTC in humans.

Author note We thank Hideki Okabayashi, Mikiya Hayashi, Masanori Kono, and Tomo Ohashi, who were kind enough to take time to comment on our draft article.

References

- Alessandri, J., Darcheville, J.-C., Delevoeye-Turrell, Y., & Zentall, T. R. (2008). Preference for rewards that follow greater effort and greater delay. *Learning & Behavior*, *36*, 352–358. doi:10.3758/LB.36.4.352
- Alessandri, J., Darcheville, J.-C., & Zentall, T. R. (2008). Cognitive dissonance in children: Justification of effort or contrast? *Psychonomic Bulletin & Review*, *15*, 673–677. doi:10.3758/PBR.15.3.673
- Amsel, A. (1958). The role of frustrative nonreward in noncontinuous reward situations. *Psychological Bulletin*, *55*, 102–119. doi:10.1037/h0043125
- Aranes, J., & Grace, R. C. (2008). Failure to obtain value enhancement by within-trial contrast in simultaneous and successive discriminations. *Learning & Behavior*, *36*, 1–11. doi:10.3758/LB.36.1.1
- Aronson, J. (1969). The theory of cognitive dissonance: A current perspective. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 4, pp. 1–34). New York: Academic Press.
- Aronson, E., & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, *59*, 177–181. doi:10.1037/h0047195
- Aw, J. M., Vasconcelos, M., & Kacelnik, A. (2011). How costs affect preferences: Experiments on state dependence, hedonic state and within-trial contrast in starlings. *Animal Behaviour*, *81*, 1117–1128. doi:10.1016/j.anbehav.2011.02.015
- Clement, T. S., Feltus, J., Kaiser, D. H., & Zentall, T. R. (2000). “Work ethic” in pigeons: Reward value is directly related to the effort or time required to obtain the reward. *Psychonomic Bulletin & Review*, *7*, 100–106. doi:10.3758/BF03210727
- Clement, T. S., & Zentall, T. R. (2002). Second-order contrast based on the expectation of effort and reinforcement. *Journal of Experimental Psychology: Animal Behavior Processes*, *28*, 64–74. doi:10.1037/0097-7403.28.1.64
- Conrad, D. G., Sidman, M., & Herrnstein, R. J. (1958). The effects of deprivation upon temporally spaced responding. *Journal of the Experimental Analysis of Behavior*, *1*, 59–65. doi:10.1901/jwab.1958.1-59
- DiGian, K. A., Friedrich, A. M., & Zentall, T. R. (2004). Discriminative stimuli that follow a delay have added value for pigeons. *Psychonomic Bulletin & Review*, *11*, 889–895. doi:10.3758/BF03196717
- Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. New York, NY: Appleton-Century-Croft.
- Friedrich, A. M., & Zentall, T. R. (2004). Pigeons shift their preference toward locations of food that take more effort to obtain. *Behavioural Processes*, *67*, 405–415. doi:10.1016/j.beproc.2004.07.001
- Gipson, C. D., Miller, H. C., Alessandri, J., & Zentall, T. R. (2009). Within-trial contrast: The effect of probability of reinforcement in training. *Behavioural Processes*, *82*, 126–132. doi:10.1016/j.beproc.2009.05.006
- Kacelnik, A., & Marsh, B. (2002). Cost can increase preference in starlings. *Animal Behaviour*, *63*, 245–250. doi:10.1006/anbe.2001.1900
- Kish, G. B. (1966). Studies of sensory reinforcement. In W. K. Honig (Ed.), *Operant behavior: Areas of research and application* (pp. 109–159). Upper Saddle River, NJ: Prentice-Hall.
- Klein, E. D., Bhatt, R. S., & Zentall, T. R. (2005). Contrast and the justification of effort. *Psychonomic Bulletin & Review*, *12*, 335–339. doi:10.3758/BF03196381
- Laties, V. G., Weiss, B., Clark, R. L., & Reynolds, M. D. (1965). Overt “mediating” behavior during temporally spaced responding. *Journal of the Experimental Analysis of Behavior*, *8*, 107–116. doi:10.1901/jeab.1965.8-107
- Melges, F. T., & Poppen, R. L. (1976). Expectation of rewards and emotional behavior in monkeys. *Journal of Psychiatric Research*, *13*, 11–21. doi:10.1016/0022-3956(76)90005-4
- Shibasaki, M., & Kawai, N. (2008). The effects of response cost and time on choosing a stimulus. *Japanese Journal of Psychology*, *79*, 241–249. doi:10.4992/jjpsy.79.241
- Shibasaki, M., & Kawai, N. (2011). The reversed work-ethic effect: Monkeys avoid stimuli associated with high-effort. *Japanese Psychological Research*, *53*, 77–85. doi:10.1111/j.1468-5884.2010.00449.x
- Vasconcelos, M., & Urcuioli, P. J. (2008). Deprivation level and choice in pigeons: A test of within-trial contrast. *Learning & Behavior*, *36*, 12–18. doi:10.3758/LB.36.1.12
- Vasconcelos, M., Urcuioli, P. J., & Lionello-DeNolf, K. M. (2007). Failure to replicate the “work ethic” effect in pigeons. *Journal of the Experimental Analysis of Behavior*, *87*, 383–399. doi:10.1901/jeab.2007.68-06
- Zentall, T. R. (2010). Justification of effort by humans and pigeons: Cognitive dissonance or contrast? *Current Directions in Psychological Science*, *19*, 296–300. doi:10.1177/0963721410383381
- Zentall, T. R. (2013). Animals prefer reinforcement that follows greater effort: Justification of effort or within-trial contrast? *Comparative Cognition and Behavior Reviews*, *8*, 60–77. doi:10.3819/ccbr.2013.80004
- Zentall, T. R., & Singer, R. A. (2007). Within-trial contrast: Pigeons prefer conditioned reinforcers that follow a relatively more rather than a less aversive event. *Journal of the Experimental Analysis of Behavior*, *88*, 131–149. doi:10.1901/jeab.2007.27-06
- Zuriff, G. E. (1969). Collateral responding during differential reinforcement of low rates. *Journal of the Experimental Analysis of Behavior*, *12*, 971–976. doi:10.1901/jeab.1969.12-971